

FetalSense: Anatomy-Guided Fetal–Probe Motion Disentanglement in Freehand Ultrasound

Sowjanya Balaji^{1*}, Anusha A^{1*}, Keerthi Ram³, Shyam Ayyasamy², Suresh Seshadri⁴, Manojkumar Lakshmanan², and Mohanasankar Sivaprakasam¹

¹ Department of Electrical Engineering, Indian Institute of Technology Madras (IITM), India

² Healthcare Technology Innovation Centre (HTIC), IITM, India

³ Sudha Gopalakrishnan Brain Centre, IITM, India

⁴ MediScan Systems, Chennai, India

ee25s080@smail.iitm.ac.in, ee25s083@smail.iitm.ac.in

Abstract. Fetal movement is an important indicator of fetal well-being and underpins routine obstetric ultrasound (US) tasks such as biometric assessment and standard-plane acquisition. Accurate motion estimation is therefore essential for both clinical analysis and autonomous US systems. However, anatomical changes observed in freehand US image sequences arise from the combined effects of fetal and probe motion, creating a motion ambiguity that makes fetal-motion estimation inherently ill-posed. To address this challenge, we introduce *FetalSense*, a framework for disentangling fetal and probe motion from US image pairs. FetalSense is trained using simulation-derived SE(3) pose supervision for both the fetus and probe, while requiring only US images and tracked probe motion at inference. We further introduce *Anatomical Landmark Displacements* (ALDs), an image-derived signal of fetal motion based on temporal changes in fetal anatomy, thereby providing supervision without requiring fetal-pose measurements at deployment. FetalSense combines a probe-pose-conditioned image encoder with an ALD-guided landmark stream to enable explicit motion decomposition. We evaluate the method on a large-scale fetal US motion dataset generated using our simulation framework, which will be released publicly upon acceptance. FetalSense achieves a mean fetal displacement error of **1.56 mm** and a decomposition closure error of **0.69 mm**. The learned fetal-motion representation further supports motion-state classification, achieving **94.1%** accuracy and **96.7%** F1-score for significant fetal-movement detection. By jointly introducing fetal–probe motion disentanglement and ALDs as an image-derived motion cue, FetalSense establishes a foundation for motion-aware fetal US analysis and autonomous US navigation.

Keywords: Fetal Ultrasound · Motion Estimation · Fetal motion flagging · Anatomical Landmark Displacements

* These authors contributed equally to this work.

1 Introduction

Fetal movement is a key indicator of neurological integrity and fetal well-being [16], detectable from as early as the second trimester [1]. Accurate characterization of fetal motion underpins a range of clinical tasks, including biometric assessment [8], standard-plane identification [2], and automated fetal monitoring [17]. In freehand ultrasound (US), however, observed anatomical displacement arises from the superposition of two independent motion sources: fetal movement within the uterus and probe movement across the maternal abdomen [5]. Because these contributions are inseparable at the image level, pixel displacement alone cannot determine whether a change is caused by fetal motion, probe motion, or both [21]. While probe motion is externally measurable via tracking, fetal motion has no analogous sensor and must be inferred from the image itself. This *motion ambiguity* directly affects clinical interpretation and downstream analysis: fetal motion can introduce imaging artifacts that hinder reliable standard-plane acquisition [14], while also confounding tasks such as standard-plane classification and fetal motion-state detection [3,13,20].

Prior work has addressed probe and fetal motion largely in isolation. On the probe side, methods such as US-GuideNet [5] and USPoint [21] estimate probe displacement from image sequences to support navigation towards target standard planes, while sensorless reconstruction approaches recover probe trajectories for volumetric compounding [15]. These methods typically assume stationary anatomy and do not account for fetal motion. Conversely, fetal motion studies have focused on landmark tracking [11,6], pose estimation [18,4], and, more recently, self-supervised motion-state classification [7]. While effective at characterizing apparent anatomical motion, these approaches implicitly attribute observed displacement to fetal activity, leaving probe-induced motion entangled with the estimated signal. Although self-supervised disentanglement methods have shown promise for separating coupled motion sources [10], the dual-source motion structure of freehand US has yet to be explicitly exploited.

We present **FetalSense**, a framework for decomposing probe-induced and fetal-induced motion from freehand ultrasound image pairs. FetalSense is trained using simulation-derived fetal and probe $SE(3)$ supervision, while requiring only ultrasound images and tracked probe motion at inference. To isolate fetal motion, we introduce *Anatomical Landmark Displacements* (ALDs), image-derived residual landmark shifts obtained by comparing observed anatomy with that expected under probe motion alone. ALDs provide a fetal-specific motion cue without requiring fetal pose measurements at deployment. By combining a probe-pose-conditioned image encoder with an ALD-guided landmark stream, FetalSense explicitly decomposes probe and fetal $SE(3)$ transformations from freehand ultrasound sequences. Our contributions are four-fold:

- **FetalSense: a framework for motion decomposition.** We present the framework to disentangle fetal and probe motion in freehand ultrasound, enabling independent estimation of fetal activity without manual fetal motion labels.

- **Anatomical Landmark Displacements (ALDs).** We introduce ALDs, residual landmark displacements that isolate fetal-specific motion by accounting for probe-induced anatomical changes. ALDs provide an anatomy-aware supervisory signal for learning fetal motion representations.
- **An open-source motion-decomposed ultrasound dataset.** We release a dataset of simulated freehand ultrasound trajectories with paired fetal and probe motion annotations.
- **Downstream fetal motion-state classification.** We demonstrate that the disentangled fetal representation supports reliable classification of static versus active fetal states, achieving 94.1% accuracy and 96.7% F1-score.

2 Dataset and Methodology

Dataset FetalSense is trained on a simulated freehand fetal ultrasound dataset derived from real patient 3-D fetal head volumes acquired at a tertiary clinical centre. The dataset comprises singleton pregnancies between 19–21 gestational weeks, a developmental window that provides both optimal fetal head visibility for standard-plane assessment and sufficient fetal motion for analysis. Free-hand scanning is simulated by generating probe trajectories through each volume and extracting 2-D ultrasound slices at every timestep. Independent fetal motion is introduced through clinically plausible rotations of the volume reference frame. For each consecutive frame pair (I_i, I_{i+1}) , ground-truth labels are available for the probe-relative transform $T_{\text{probe,rel}}^{(i)}$, fetal-relative transform $T_{\text{fetus,rel}}^{(i)}$, and combined transform $T_{\text{total,rel}}^{(i)}$, together with a binary fetal motion-state label. Beyond SE(3) pose labels, each image pair contains Anatomical Landmark Displacements (ALDs) and per-landmark visibility masks. Landmark centroids extracted from the segmentation module are used to compute ALDs, which quantify fetal-induced motion as the displacement between observed and probe-predicted centroid locations: $\mathbf{a}_k^{(i)} = \mathbf{c}_{k,\text{obs}}^{(i+1)} - \mathbf{c}_{k,\text{pred}}^{(i+1)}$, where $\mathbf{c}_{k,\text{obs}}^{(i+1)}$ and $\mathbf{c}_{k,\text{pred}}^{(i+1)}$ denote the observed and probe-predicted centroid locations of landmark k . By accounting for probe-induced motion, ALDs provide localized, anatomy-aware indicators of fetal movement. Each ALD is accompanied by a visibility mask $v_k^{(i)} \in \{0, 1\}$, with supervision applied only to visible landmarks. The rendered 2-D slices retain authentic ultrasound characteristics, including speckle, shadowing, and tissue contrast, while speckle variability from 0.25 to 0.57 and SNR variability from 3.11 to 9.18 across 55 patient volumes expose the model to diverse clinical imaging conditions during training.

FetalSense Architecture For consecutive frames (I_i, I_{i+1}) , the observed image motion is jointly determined by probe motion and fetal motion. Let $T_{\text{probe,rel}}^{(i)} = \left(T_{\text{probe}}^{(i)}\right)^{-1} T_{\text{probe}}^{(i+1)}$ and $T_{\text{fetus,rel}}^{(i)} = \left(T_{\text{vol}}^{(i)}\right)^{-1} T_{\text{vol}}^{(i+1)}$ denote the relative probe and fetal transforms between timesteps i and $i + 1$. Expressing the probe pose in the fetal reference frame yields $T_{\text{in-vol}}^{(i)} = \left(T_{\text{vol}}^{(i)}\right)^{-1} T_{\text{probe}}^{(i)}$ and

$T_{\text{total,rel}}^{(i)} = \left(T_{\text{in-vol}}^{(i)}\right)^{-1} T_{\text{in-vol}}^{(i+1)}$, where $T_{\text{total,rel}}^{(i)}$ represents the observed inter-frame motion. Although probe motion is directly available from tracking, estimating fetal motion remains challenging because the observed image motion reflects the combined effects of both probe and fetal movement. To address this challenge, FetalSense employs a dual-input architecture comprising a FiLM-conditioned EfficientNet image encoder modulated by tracked probe motion and a visibility-masked landmark encoder that processes ALDs and visibility masks, where ALDs provide localized cues of fetal movement. Features from both streams are fused and processed by dedicated total-motion, fetal-motion, and ALD prediction branches. The total-motion and fetal-motion branches estimate the observed and fetal motion components, while the ALD branch provides anatomy-aware supervision that links fetal motion predictions to observed landmark displacements. Joint optimization of all three branches enables accurate and geometrically consistent motion estimation.

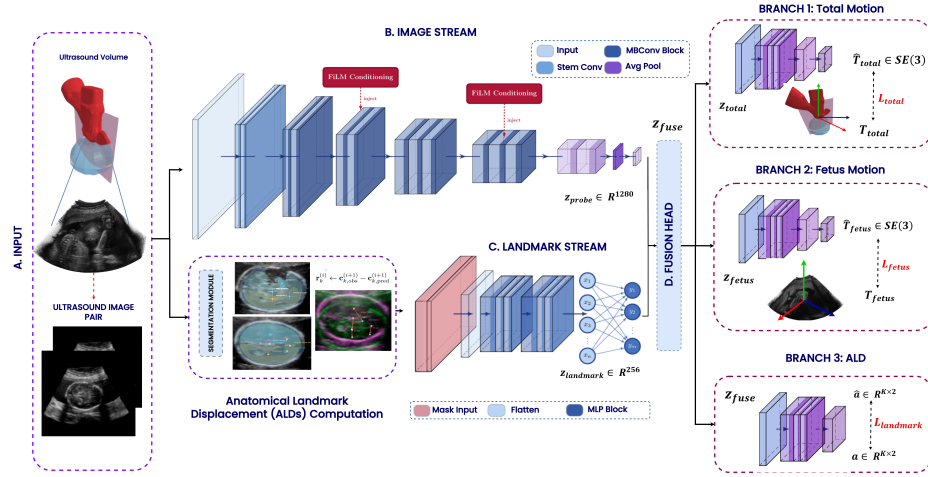


Fig. 1. Overview of FetalSense. (A) **Input.** Consecutive ultrasound frames (I_i, I_{i+1}) with probe-, fetal-, and total-motion labels, and Anatomical Landmark Displacements (ALDs). (B) **Image Stream.** A FiLM-conditioned EfficientNet extracts motion features from image pairs conditioned on probe motion. (C) **Landmark Stream.** A visibility-masked landmark encoder extracts anatomy-aware fetal-motion features from ALDs. (D) **Fusion and Prediction Branches.** Features are fused and decoded into total-motion, fetal-motion, and ALD predictions.

Image Stream with Probe Conditioning Given a consecutive ultrasound image pair (I_i, I_{i+1}), the two frames are stacked and processed by an EfficientNet-B1 encoder to produce a visual representation z_{probe} . Since the observed image motion reflects both probe and fetal movement, FetalSense explicitly incorporates tracked probe motion through Feature-wise Linear Modulation (FiLM). Let $\mathbf{p}_i$ denote the probe-motion representation derived from $T_{\text{probe,rel}}^{(i)}$. A conditioning network predicts feature-wise modulation parameters (γ_i, β_i) , which modulate intermediate feature maps according to

$$\tilde{\mathbf{z}}_{\text{probe}} = \gamma_i \odot \mathbf{z}_{\text{probe}} + \beta_i. \quad (1)$$

Unlike conventional image encoders, FiLM conditioning injects known probe motion directly into the feature hierarchy. FiLM layers are applied at multiple intermediate stages of EfficientNet-B1, allowing probe motion to influence both low-level texture cues and higher-level anatomical representations. The resulting probe-aware visual embedding $\mathbf{z}_{\text{probe}} \in \mathbb{R}^{1280}$ encodes appearance and motion information conditioned on tracked probe displacement.

Landmark Stream The landmark stream encodes Anatomical Landmark Displacements (ALDs) and visibility masks derived from fetal head landmarks. Since ALDs capture landmark motion unexplained by tracked probe movement, they provide localized indicators of fetal motion. Visibility-masked ALDs are flattened and encoded by a three-layer MLP into a compact anatomy-aware embedding $\mathbf{z}_{\text{landmark}} = h_\psi([\mathbf{a}_1, \dots, \mathbf{a}_K] \odot \mathbf{v})$, where $\mathbf{v}$ denotes the landmark visibility mask and h_ψ is a three-layer MLP with LayerNorm and GELU activations. The resulting embedding $\mathbf{z}_{\text{landmark}} \in \mathbb{R}^{256}$ captures spatially coherent patterns of fetal motion across anatomically meaningful structures.

Fusion Head and Prediction Branches The probe-conditioned image embedding $\mathbf{z}_{\text{probe}} \in \mathbb{R}^{1280}$ and landmark embedding $\mathbf{z}_{\text{landmark}} \in \mathbb{R}^{256}$ are concatenated and projected into a shared latent representation $\Phi([\mathbf{z}_{\text{probe}}, \mathbf{z}_{\text{landmark}}])$, where $\Phi(\cdot)$ denotes the fusion trunk. The resulting representation combines complementary visual and anatomy-aware motion cues extracted from the image and landmark streams. To encourage specialization between global and fetal motion estimation, the shared representation is processed by two private branches, which are used to predict total motion, fetal motion, and the ALD field:

$$\mathbf{z}_{\text{total}} = B_{\text{total}}(\mathbf{z}_{\text{fuse}}), \quad \mathbf{z}_{\text{fetus}} = B_{\text{fetus}}(\mathbf{z}_{\text{fuse}}), \quad (2)$$

$$\hat{T}_{\text{total,rel}}^{(i)} = H_{\text{total}}(\mathbf{z}_{\text{total}}), \quad \hat{T}_{\text{fetus,rel}}^{(i)} = H_{\text{fetus}}(\mathbf{z}_{\text{fetus}}), \quad \hat{\mathbf{a}} = H_{\text{Landmark}}(\mathbf{z}_{\text{fuse}}). \quad (3)$$

Here, $\hat{T}_{\text{total,rel}}^{(i)}$ and $\hat{T}_{\text{fetus,rel}}^{(i)}$ denote the predicted total and fetal relative motions, respectively, while $\hat{\mathbf{a}}$ denotes the predicted ALD field. The ALD branch provides localized anatomy-aware supervision that guides fetal motion estimation through consistency with observed landmark displacements, whereas the private motion branches encourage the total-motion and fetal-motion heads to learn distinct representations. All branches are optimized jointly using

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{total}} + \lambda_2 \mathcal{L}_{\text{fetus}} + \lambda_3 \mathcal{L}_{\text{Landmark}}, \quad (4)$$

where $\mathcal{L}_{\text{total}}$ and $\mathcal{L}_{\text{fetus}}$ supervise total and fetal pose estimation, respectively, and $\mathcal{L}_{\text{Landmark}}$ enforces consistency with the observed anatomical landmark displacements. While $\mathcal{L}_{\text{total}}$ and $\mathcal{L}_{\text{fetus}}$ rely on simulation-only ground-truth poses during training, inference requires only the image pair and tracked probe motion. This absence of fetal-pose access at inference is what allows the trained model to generalize to real freehand sequences where fetal motion cannot be measured.

3 Results and Discussion

Implementation Details The FetalSense dataset comprises 95,000 pairwise fetal US samples, each consisting of two consecutive 224×224 frames with corresponding probe-relative, fetal-relative, and total relative pose annotations. Data were split at the volume level into training, validation, and test sets (60:20:20). FetalSense was trained for 1000 epochs using Adam (learning rate 1×10^{-4} , batch size 32). The training objective combines total-motion, fetal-motion, and ALD losses with weights $\lambda_{\text{total}} = 1.0$, $\lambda_{\text{fetus}} = 3.0$, and $\lambda_{\text{Landmark}} = 0.1$, selected empirically on the validation set to balance motion estimation and anatomy-aware supervision. All experiments were conducted on an NVIDIA GeForce RTX 4090 GPU, achieving an average inference latency of 0.47 ms per frame pair.

Motion Estimation Performance Table 1 compares FetalSense against four baselines using fetal error (F.E.), closure error (C.E.), and accuracy. All trainable methods were trained using three independent random seeds, and the reported values are averaged across runs. Closure error measures decomposition consistency, $\|\hat{p}_{\text{fetus}} + p_{\text{probe}}^{\text{GT}} - p_{\text{total}}^{\text{GT}}\|$, with lower values indicating more accurate reconstruction of total motion, while Accuracy (%) measures moving/non-moving fetal-state classification performance. Image Stream Only performs worst overall, showing that image and probe cues alone are insufficient to separate fetal and probe motion. LightGlue [9] improves classification accuracy through correspondence-based probe compensation but exhibits high fetal error, while CURL [7], adapted from a contrastive fetal-movement detector, substantially reduces both fetal and closure errors, achieves the strongest baseline performance. We further evaluate a RAFT [19]-based sensorless baseline that estimates total motion via optical flow and subtracts known probe motion. Despite using the same probe information as FetalSense, RAFT achieves reasonable classification accuracy of 70.65% but substantially higher F.E and C.E. In contrast, FetalSense achieves the lowest F.E. lowest C.E., and highest accuracy (**94.1%**), demonstrating that landmark supervision enables more accurate and internally consistent fetal-motion estimation.

Table 1. Comparison with baseline methods.

Method	F.E (mm)↓	C.E (mm)↓	Accuracy (%)↑
Image Stream only	3.81	7.89	27.86
Lightglue [9]	5.16	6.31	56.00
CURL (fine-tuned) [7]	2.95	3.69	63.38
RAFT + Probe Subtraction [19]	4.19	4.33	70.65
FetalSense (Ours)	1.56	0.69	94.12

Ablation Study Table 2 ablates FetalSense by progressively adding the landmark stream, private pose heads, FiLM conditioning, and the ALD loss to an image-stream baseline. The landmark stream alone yields modest gains, as both motion heads still share backbone features. Private pose heads markedly reduce

fetal and closure error, showing that task-specific representations are needed to disentangle probe- from fetal-induced motion. FiLM conditioning further lowers fetal error by injecting probe geometry directly into the image stream, but at the cost of closure consistency, which is resolved only once the ALD loss is added. The full model achieves the best fetal error (**1.56 mm**), closure error (**0.69 mm**), and cosine similarity (-0.15), indicating reduced alignment between the total- and fetal-motion predictions. Together, these results suggest more effective disentanglement of probe- and fetal-induced motion.

Table 2. Ablation study of FetalSense.

Configuration	F.E (mm)	T.E (mm)	C.E (mm)	Cosine Sim.
Image Stream Only	3.81 ± 2.8	2.7 ± 1.3	7.89 ± 5.9	-0.19
Dual Stream+Shared Branches	3.35 ± 2.7	3.08 ± 1.6	7.58 ± 6.3	-0.21
Dual Stream+Private Branches	1.95 ± 1.9	2.46 ± 1.6	4.27 ± 4.1	-0.16
Dual Stream+Private Branches+FiLM	1.67 ± 1.1	1.75 ± 1.7	4.42 ± 4.6	-0.23
FetalSense (Ours)	1.56 ± 1.8	1.77 ± 1.7	0.69 ± 1.4	-0.15

Motion Representation Analysis A key objective of FetalSense is to disentangle fetus-specific motion from global probe motion rather than learn a single mixed representation. We therefore evaluate whether the total-motion and fetal-motion branches learn distinct and accurate representations (Fig. 2). Panels (a–b) compare predicted and ground-truth total and fetal motion, while (c) assesses decomposition consistency. Panel (d) examines still-fetus intervals, where fetal-motion predictions should remain near zero, and (e) evaluates branch redundancy via cosine similarity; quantitative results are summarized in (f). Median reconstruction errors are 2.18 mm and 0.84 mm for total and fetal motion, respectively. Strong within-branch correlations and weaker cross-correlations indicate reduced coupling between the motion branches, while a low ALD prediction error (0.39 px) suggests accurate capture of localized fetus-specific motion. Together, these results demonstrate accurate motion decomposition and complementary motion representations.

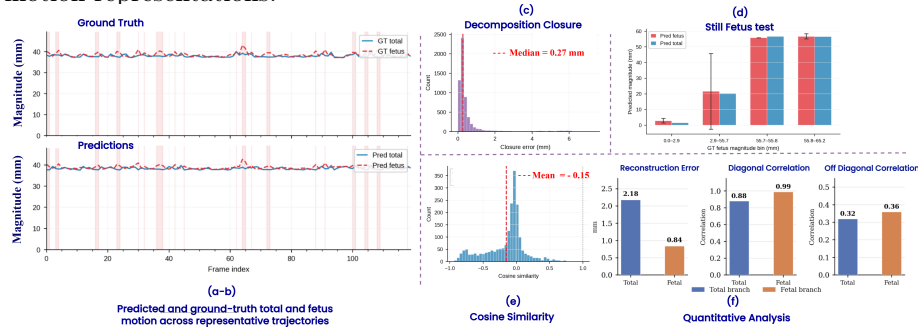


Fig. 2. Motion decomposition analysis. (a–b) Total and fetal motion predictions versus ground truth. (c) Decomposition consistency. (d) Still-fetus predictions. (e) Inter-branch cosine similarity. (f) Reconstruction and correlation metrics.

Downstream Motion-State Classification Identifying periods of fetal activity is important for autonomous US navigation [12], where probe-control decisions may be adapted during substantial fetal motion. We therefore evaluate whether the learned fetal-motion representation supports motion-state classification, categorizing frame transitions as *moving* or *static*. Figure 3 shows representative high-, low-, and mid-motion trajectories. Motion events are consistently detected across all trajectories, with no observed false negatives. Most errors occur near moving–static transition boundaries, where elevated motion variance occasionally produces false positives. High- and mid-motion trajectories achieve F1 scores of 98.8% and 97.3%, respectively, while the low-motion trajectory is more challenging due to class imbalance. These results indicate that the learned representation captures clinically meaningful fetal-motion dynamics beyond pose estimation alone.

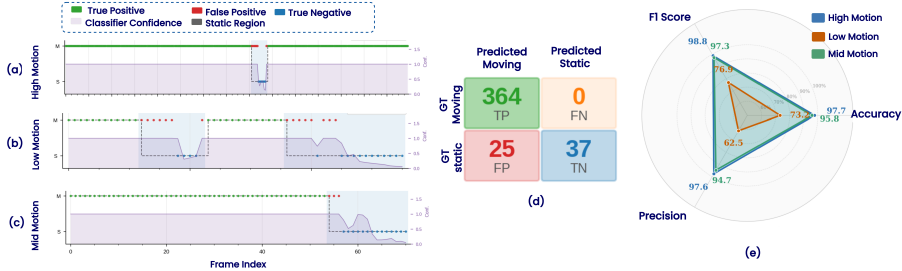


Fig. 3. Downstream motion-state classification. (a–c) Ground-truth and predicted states across representative trajectories. (d) Aggregate confusion matrix. (e) Per-trajectory classification performance.

4 Conclusion and Future Works

FetalSense presents a framework for disentangling fetal and probe-induced motion in freehand ultrasound using Anatomical Landmark Displacements (ALDs) as geometry-aware motion cues. By jointly leveraging image appearance and anatomy-aware landmark dynamics, the proposed method achieves accurate and geometrically consistent motion decomposition, with a mean fetal displacement error of 1.56 mm and a closure error of 0.69 mm. Beyond motion estimation, the learned fetal representation supports motion-state classification, achieving 94.1% accuracy and a 96.7% F1-score. These results demonstrate the potential of anatomy-guided supervision for fetal motion understanding and provide a foundation for motion-aware fetal ultrasound analysis and autonomous ultrasound navigation. Future work will focus on bridging the simulation-to-real gap through validation on real freehand fetal ultrasound scans, benchmarking the disentangled fetal representation in downstream freehand reconstruction pipelines, and extending the framework to broader fetal anatomies and closed-loop autonomous scanning.

References

1. Fetal movement. In *ScienceDirect Topics — Medicine and Dentistry*. Elsevier. Available at: <https://www.sciencedirect.com/topics/medicine-and-dentistry/fetal-movement>.
2. Christian F. Baumgartner, Konstantinos Kamnitsas, Jacqueline Matthew, Tim P. Fletcher, Sandra Smith, Lisa M. Koch, Bernhard Kainz, and Daniel Rueckert. SonoNet: Real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Transactions on Medical Imaging*, 36(11):2204–2215, 2017.
3. J. Brändle, H. Preissl, R. Draganova, E. Ortiz, K. O. Kagan, H. Abele, S. Y. Brucker, and I. Kiefer-Schmidt. Heart rate variability parameters and fetal movement complement fetal behavioral states detection via magnetography to monitor neurovegetative development. *Frontiers in Human Neuroscience*, 9:147, 2015.
4. Chaoyu Chen, Xin Yang, Yuhao Huang, Wenlong Shi, Yan Cao, Mingyuan Luo, Xindi Hu, Lei Zhu, Lequan Yu, Kejuan Yue, Yuanji Zhang, Yi Xiong, Dong Ni, and Weijun Huang. Fetusmapv2: Enhanced fetal pose estimation in 3d ultrasound. *Medical Image Analysis*, 91:103013, 10 2023.
5. Richard Droste, Lior Drukker, Aris T. Papageorgiou, and J. Alison Noble. Automatic probe movement guidance for freehand obstetric ultrasound. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 12263 of *Lecture Notes in Computer Science*, pages 583–592. Springer, Cham, 2020.
6. Hengtao Guo, Sheng Xu, Bradford Wood, and Pingkun Yan. Sensorless freehand 3d ultrasound reconstruction via deep contextual learning. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III*, Berlin, Heidelberg, 2020. Springer-Verlag.
7. X. He et al. CURL: Contrastive unsupervised representation learning for fetal motion. *IEEE Transactions on Medical Imaging*, 44(1):1–12, 2025.
8. Chihiro Komatsu et al. Automatic identification of fetal abdominal planes from ultrasound images based on deep learning. *PMC (PubMed Central)*, 2025. PMC article ID: PMC12701097. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12701097/>.
9. Philipp Lindenberger, Paul-Edouard Sarlin, and Marc Pollefeys. Lightglue: Local feature matching at light speed, 2023.
10. Xiao Liu, Pedro Sanchez, Spyridon Thermos, Alison Q. O’Neil, and Sotirios A. Tsaftaris. Learning disentangled representations in the imaging domain. *Medical Image Analysis*, 80:102516, 2022.
11. Chen Ma, Yunshu Li, Bowen Guo, Jing Jiao, Yi Huang, Yuanyuan Wang, and Yi Guo. Unlabeled data-driven fetal landmark detection in intrapartum ultrasound. In Jieyun Bai, Yuxin Huang, Isaac Khobo, Mohammad Yaqub, Karim Lekadir, Dong Ni, and Shuo Li, editors, *Intrapartum Ultrasound*, pages 14–23, Cham, 2026. Springer Nature Switzerland.
12. F. A. Manning, T. F. Baskett, I. Morrison, and I. Lange. Fetal biophysical profile scoring: A prospective study in 1,184 high-risk patients. *Am. J. Obstet. Gynecol.*, 140(3):289–294, 1981.
13. J. G. Nijhuis, H. F. R. Prechtl, C. B. Martin, and R. S. G. M. Bots. Are there behavioural states in the human fetus? *Early Human Development*, 6(2):177–195, 1982.

14. Niroshan Perera et al. Automated segmentation of standard scanning planes to measure biometric parameters in fetal ultrasound images — a survey. *Research-Gate preprint*, 2023. Available at: <https://www.researchgate.net/publication/368350412>.
15. Raphael Prevost, Mehrdad Salehi, Simon Jagoda, Navneet Kumar, Julian Sprung, Alexander Ladikos, Robert Bauer, Oliver Zetting, and Wolfgang Wein. 3d free-hand ultrasound without external tracking using deep learning. In *Medical Image Analysis*, volume 48, pages 187–202, 2018.
16. Nandini Raghuraman and Leela Sharath Pillarisetty. Fetal movement. StatPearls [Internet]. StatPearls Publishing, Treasure Island (FL), 2026. Updated March 25, 2026. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK470566/>.
17. Bahbib Rahmatullah, Aris T. Papageorghiou, and J. Alison Noble. Automated scoring of fetal abdomen ultrasound scan-planes for biometry. In *Proceedings of the International Symposium on Biomedical Imaging (ISBI)*, 2015. Available at: <https://www.researchgate.net/publication/275155742>.
18. Jayroop Ramesh, Nicola Dinsdale, Pak-Hei Yeung, and Ana I. L. Namburete. Geometric Transformation Uncertainty for Improving 3D Fetal Brain Pose Prediction from Freehand 2D Ultrasound Videos . In *proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, volume LNCS 15001. Springer Nature Switzerland, October 2024.
19. Zachary Teed and Jia Deng. Raft: Recurrent all-pairs field transforms for optical flow, 2020.
20. Judith O. E. H. van Laar et al. Evaluation of parameters for fetal behavioural state classification. *Scientific Reports*, 12:3498, 2022.
21. Xin Yang et al. USPoint: Self-supervised interest point detection and description for ultrasound-probe motion estimation during fine-adjustment standard fetal plane finding. *PMC (PubMed Central)*, 2020. PMC article ID: PMC7614558. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC7614558/>.